THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 50, No. 4, April, 1971
Printed in U.S.A.

1969-70 Connection Survey:

Low-Speed Data Transmission Performance on the Switched Telecommunications Network

By H. C. FLEMING and R. M. HUTCHINSON, JR.

(Manuscript received December 4, 1970)

Error statistics of low-speed, start-stop data transmission over the Bell System switched telecommunications network are reported in this article. During the 1969–70 Toll Connection Survey, measurements were made on 534 connections with over 21 million characters transmitted, to give an overall average character error rate of 1.46 × 10^{-4}. Over 90 percent of the low-speed test calls contained about 36,000 or 54,000 characters. A character error rate of 10^{-4} or less is indicated for 77.6 percent of all calls, while 95.0 percent of calls have a lost character rate of 10^{-4} or less. Error-free performance is shown for 48.3 percent of all calls, and 89.3 percent have no lost characters. Statistics of bursts, error-free intervals, block errors, and carrier failure durations are also presented.

I. INTRODUCTION

1.1 *Reasons for Measuring Low-Speed Data Performance*

Measurements of the accuracy of low-speed data transmission over the Bell System switched network have been included in the 1969–70 Toll Connection Survey; the results are presented in this article. There is widespread use of teletypewriters, computer ports, and other terminal devices that communicate by means of data organized in characters comprised of several bits, using start-stop transmission, at rates up to 300 bits per second (b/s). About 80,000 terminals of this type were being used for *Data-Phone*® service at the end of 1970.

This article presents the quantity and distribution of the character errors and lost characters observed during the low-speed data tests. These results are expected to meet a need for characterization of the switched telecommunications network for start-stop transmission of

low-speed data. Further analysis of the data is planned to pinpoint causes of errors for which remedial measures would be profitable, to suggest possible improvements in design of transmission equipment, and to indicate possible usefulness of error control systems.

The measurements were planned with the objective of characterizing low-speed data transmission performance as viewed by users. The error statistics are reported for channels extending from a transmitting data set, into which an error-free data signal is sent, to the point at the receiving end at which the characters are used or displayed.

1.2 *Design of Survey*

Detailed planning of the survey is described in a companion article by F. P. Duffy and T. W. Thatcher[1]; a very brief summary is presented here. A random selection of end points for approximately 600 toll calls was made by means of a three-stage sampling plan. The first stage consisted of division of the 48 contiguous United States and the provinces of Ontario and Quebec in Canada into 12 geographical areas that originated roughly equal numbers of completed toll calls. A "primary" end office was randomly chosen from each area with probability proportional to the originated toll traffic from that office. In the second stage, two to six calls which determined "secondary" or far-end test offices were selected randomly in each of three mileage bands (0–180, 180–725, and 725–2900 miles) relative to each primary office location. The calls were selected from traffic printouts from each primary office; thus the chosen secondary offices had probability of selection proportional to traffic originated from the primary office destined to the secondary office. The third stage consisted of six or more test calls dialed from each primary office to each of its associated secondary offices. Numerous analog transmission parameters and error performance for four Bell System data sets were measured on most connections.[1] Results for the Data Sets 201(2000 b/s), 202(1200 b/s), and 203(3600 and 4800 b/s) are reported by M. D. Balkovic, H. W. Klancer, S. W. Klare, and W. G. McGruther.[2] This article gives results for Data Set 103 at 150 b/s.

1.3 *Start-Stop Transmission*

The data transmitted at low speed were organized into successive 10-bit characters, and were sent in a start-stop mode. In this type of transmission the character consists of an initial "start" bit (binary 0, or space), followed by seven information bits used to encode 128 different characters—alphabetic, numeric, punctuation, and control

(American Standard Code for Information Interchange, ASCII), followed by a parity check bit (selected to give an even number of ones for the information bits and parity bit), and terminated by a "stop" interval of indefinite duration, one bit minimum (binary 1, or mark). Many terminal devices are designed to generate and respond to characters thus constituted.

Character error statistics, rather than bit error statistics, are the parameters of interest in this type of transmission because the message consists of a display (in teletypewriters) or use (in computers) of characters in most applications. Start-stop transmission is susceptible to sequences of character errors caused by a single bit error. This occurs because of the mechanism used to distinguish the start of a new character: the transition from a stop interval (mark) to a start bit (space). A disturbance that produces a space during a prolonged stop pulse, or changes a start bit (space) to a mark thus prolonging the stop pulse, will cause the receiving terminal to detect the start of the next character at the wrong point in the bit sequence and to record a character error. This condition—loss of character synchronization—persists until the correct stop-start transition is detected in a subsequent character. Some receiving circuits, for example regenerators, have built-in algorithms that minimize the resynchronizing interval. In other cases, the correct position is reached accidentally as a result of the pseudo-random bit pattern of the usual message. In short, loss of character synchronization may result from a single bit error and cause propagation of many character errors—an effect which is not present in synchronous transmission.

II. IMPLEMENTATION

2.1 *Test Arrangement*

Figure 1 shows the test arrangement used to measure the low-speed data transmission performance. The data were always transmitted from the secondary site because considerably less equipment was required at the transmitting end than at the receiving end with this arrangement, and tests were conducted for two weeks or more at a primary location compared to one or two days at a secondary site.

2.1.1 *Transmitting End*

A test sentence generator (WECo 911B) generated a repeated 72-character test sentence (FOX sentence) using ASCII 10-bit characters at 15 characters per second, or 150 b/s. The 72-character

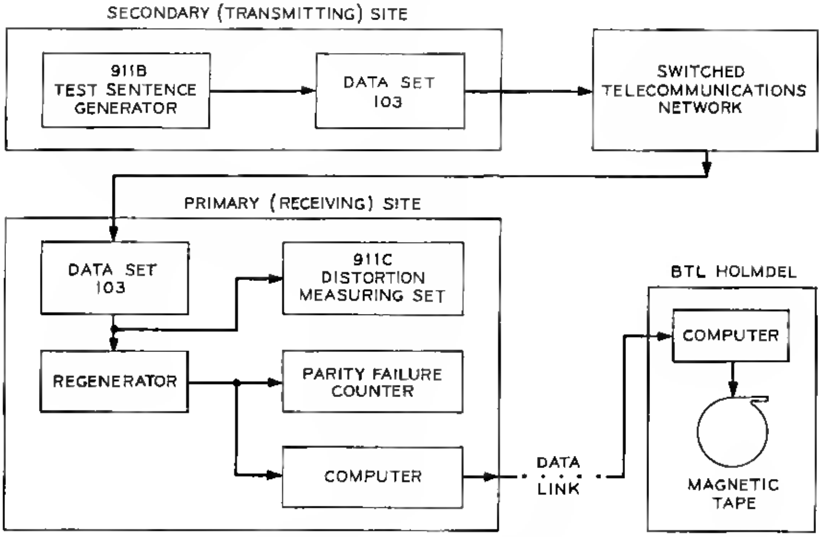

Fig. 1—150 b/s data transmission test of the 1969–70 Toll Connection Survey.

sentence produces the following line of printing on a teletypewriter followed by a delete character, two carriage returns, and a line feed: THE QUICK BROWN FOX JUMPED OVER A LAZY DOG'S BACK 1234567890 TESTING. A Bell System 103 type data set was chosen for the low-speed data transmission tests. This set is typical of the 100 series sets which are used to transmit low-speed serial data in all *Data-Phone* applications. It uses frequency shift keying (fsk) between 1070 (space) and 1270 (mark) Hz (low band) in one direction of transmission and 2025 (space) and 2225 (mark) Hz (high band) in the other direction, at rates up to 300 b/s. The 103 data set at the transmitting site modulated the test sentence to a binary fsk signal at either the low-band (F_1) or high-band (F_2) frequencies; these were alternated on successive calls.

The output signal power delivered to the connection was between -12 dBm and -13 dBm; both F_1 and F_2 were transmitted at the same power. These are the values normally produced at the central office. In a standard installation, there is a subscriber loop from the customer's premises to the main frame in the central office. In the survey, the test setup was connected to the main frame, thus including a loop of essentially zero loss. It was considered unnecessary to add real loops to customer locations because loops should not degrade the

error performance of a connection significantly. Moreover, real loops are nonrandom in nature; that is, a subscriber station uses the same loop for all calls.

2.1.2 *Receiving End*

The data set at the receiving site demodulated the binary fsk received signal to baseband mark or space voltage. The data set contains a frequency modulation discriminator which continuously makes the binary decision as to whether marking or spacing frequency is being received. There is no timing or sampling associated with this decision.

The data signal at the output of the receiving data set differs from the signal at the input to the transmitting data set because of imperfections in the transmission characteristics of the connection. One difference is displacement in time of the mark to space transitions—telegraph distortion. This was measured by the distortion measuring set (WECo 911C) in Fig. 1. Receiving terminal devices differ in their error performance susceptibility to telegraph distortion. The test arrangement included a regenerator at the output of the receiving data set which reduces the telegraph distortion to a negligible amount. Thus, the test results apply to any terminal arrangement that uses this regenerator, and are independent of the following receiving circuits. With this arrangement, the receiving terminal—regenerator plus detecting circuits—resembles many terminals in use. The specific regenerator employed is one used optionally with the Model 37 Teletypewriter.[3]

The computer at the primary site interpreted the regenerator output in terms of ASCII characters in the same manner that a teletypewriter would in most instances. The computer detected and recorded characters received in error and carrier failure indications, which occur when the received line signal level drops to a very low value (approximately -50 dBm or lower).

The parity failure counter incremented whenever a received character failed parity. Since a large fraction (about 80 percent) of characters received in error will also fail parity, the parity count provided the testers with an immediate indication of the occurrence of character errors.

2.2 *Data Collection*

Information identifying each low-speed data test run and giving measurements of parity failure counts and distortion was recorded

on log sheets at the primary location. It was also entered into the primary site computer by teletypewriter, and then transmitted over a data link to a computer at Bell Telephone Laboratories in Holmdel, New Jersey, where it was stored on magnetic tape. The data link was a switched network connection with transmission accuracy enhanced by error detection and block retransmission.

The computer at the primary site was programmed to recognize characters from the low-speed data regenerator and compare them with the expected characters, making use of a FOX sentence which was stored in its memory. Any discrepancy in the character's bits was considered to be a character error; the computer recorded the identity (bit pattern) of the received erroneous character and the number of characters received since the last previously recorded event.

As described earlier, character synchronization may be lost as the result of an error in a start or stop bit, usually causing several successive character errors. Character synchronization might be recovered on the wrong character—a condition described as loss of sentence synchronization. To prevent the computer from recording character errors interminably, when five character errors were detected in succession the computer was declared to be out of sentence synchronism (OSS) and a resynchronizing procedure was begun. During an OSS period, a complete record of the bits in all received characters was taken.

The recorded information described above and a record of the time of occurrence and duration of all carrier failures were transmitted over the data link to the Holmdel computer. This information was recorded on magnetic tape, along with the other low-speed test information mentioned above and other similar information about the high-speed data tests.[2]

2.3 *Data Processing*

The low-speed test results were separated from the other information on the magnetic tape and edited. The result was a complete record of (*i*) the bit pattern and time of occurrence of all characters, except those removed by the editing, reported by the primary site computer (isolated character errors and characters reported during OSS periods) and (*ii*) the time of beginning and end of carrier failure indications and OSS conditions.

Every transmitted character was classified as to whether it was received correctly, received in error, or not received (lost)—a conse-

quence of a carrier failure. The first character in a carrier failure was defined to be a character error, since this character is very likely to be received in error because of loss of signal during the transmission of the character. A teletypewriter would very probably print an error for this character. (This definition was necessary because the primary site computer did not check this character for bit errors.)

OSS intervals were analyzed as follows. The received characters reported by the computer were compared with the FOX sentence characters. When the received character was in the FOX sentence, checks of subsequent characters determined whether sentence sync had been recovered. This generally occurred before the computer completed its resynchronizing procedure, which was based on detection of a specific FOX sentence character. Characters transmitted during that part of an OSS interval preceding recovery of sentence sync were character errors. Status of characters following recovery of sync was based on comparison with the corresponding FOX sentence characters.

Finally, cards were punched for each call containing information identifying the call and summarizing the character statistics. In addition, the entire call was recorded on cards in a format which described the exact sequence of characters received correctly, in error, or lost. This information was condensed by recording the length (in characters) of intervals of received characters for each category.

This set of cards comprised the input data from which all error statistics reported in the remainder of this article were computed.

III. DISTRIBUTIONS OF CHARACTER ERROR AND LOST CHARACTER RATES* OF CALLS

3.1 *Calls Included in Analysis*

The set of calls for which data transmission error statistics are reported in this article is the same as that for which error statistics are reported for the high-speed voiceband data sets,[2] except for connections on which both low-speed and high-speed data measurements could not be completed. For reasons given in companion articles,[1,2] the set differs somewhat from the set selected by the sampling plan described earlier. The essential difference is that two primary offices in the original sample were replaced by nearby offices, from which high-

* Character error rate of a call = character errors/characters transmitted in that call.

speed voiceband data users in that area would normally be served. In addition, all calls that were switched by panel type central office equipment at either end of the connection were excluded; panel offices do not normally serve data customers. Only one primary office contained panel equipment, but it also contained crossbar equipment and calls originated through the latter equipment were included.

3.2 *Magnitude of Low-Speed Data Survey*

Table I summarizes the quantities of calls and primary and secondary locations.

The duration of a low-speed data test was nominally 40 minutes except when the 3600–4800 b/s data set was not tested (for about half the calls in the first half of the survey). In this case, the duration was about an hour. Equipment problems and dropped connections sometimes resulted in shorter runs—13 calls shorter than 23 minutes, with a minimum of 1.1 minutes. In 15 cases, extra available time was used for tests longer than 65 minutes, with a maximum of 145 minutes.

Table II summarizes some of the total survey statistics. Character errors are characters in which at least one bit is received in error, with the result that an erroneous character may be printed or an undesired action may take place (for example, carriage return) on the receiving teletypewriter. Lost characters are those for which no bits are received, usually because of a loss of the received signal, resulting in characters missing from the teletypewriter printout. Statistics are presented separately for these two types of impairments because they differ in the extent to which they damage the communication. Character errors may result in acceptance of incorrect information (an order for one carload might be received as an order for nine carloads), whereas lost characters represent missing information, perhaps requiring retransmission of the message.

Table I—Call and Location Statistics

	Total	Mileage Band		
		Short	Medium	Long
Calls	534	171	186	177
Primary Locations	12			
Secondary Locations	91	29	32	30
Secondaries Per Primary	5 to 11			
Calls Per Primary	22 to 62	5 to 33	6 to 29	8 to 23

TABLE II—TOTAL SURVEY CHARACTER ERROR AND LOST CHARACTER STATISTICS

	All Calls	Mileage Band Short	Medium	Long
Calls	534	171	186	177
Characters Transmitted	21.31×10^6	7.00×10^6	7.48×10^6	6.83×10^6
Character Errors	3,110	751	1,053	1,306
Lost Characters	14,511	9,581	1,476	3,454
Character Error Rate	1.46×10^{-4}	1.07×10^{-4}	1.42×10^{-4}	1.90×10^{-4}
Lost Character Rate	6.81×10^{-4}	13.7×10^{-4}	1.98×10^{-4}	5.03×10^{-4}

3.3 *Overall Average Rates*

Overall average error rates are also listed in Table II. These are useful as single number parameters that characterize the survey. Distributions of per call statistics are more significant; these are presented later in this article. The average character error rate—total character errors observed during the survey divided by total characters transmitted—is 1.46×10^{-4}, which corresponds to one error for every 1.7 pages of average single-spaced typed text. As Fig. 2, curve (A), shows, the average character error rate was not dominated by a few bad calls; it drops by a factor of only 1.6 (to 9.1×10^{-5}) when the eight worst calls are omitted.

The average lost character rate is 6.81×10^{-4}, corresponding to about 2.7 lost characters per page of text. This average is definitely dominated by a few bad calls; omission of the eight worst calls, which contain 93.3 percent of the total lost characters, improves the average by a factor of 14.5 to 1, to 4.7×10^{-5} [Fig. 2, curve (B)].

The average character error rate becomes progressively poorer as the mileage band gets longer, although the span of the increase is not large, only a factor of 1.8. Lost character rate is poorest in the shortest mileage band; however, omission of two calls—the worst and third worst for all calls—improves the rate in the short band to 1.55×10^{-4}, which is better than the medium and long bands.

3.4 *Weighting of Data for Calls*

Error statistics of calls include application of weighting factors to each call, based on the details of the sampling procedure by which the set of calls was selected. One contribution to the weighting arises

from unequal toll traffic originated within the 12 areas into which the United States and Canada were divided. A second and much larger component of the weighting factor which is applied to the set of all calls results from large differences in the total toll traffic from a primary to offices in each mileage band. A third component is the actual number of calls for which low-speed data error statistics were collected between specific primary and secondary locations. In this article, weighting was applied only to the error statistics of calls presented in Section 3.5 and not to the remaining parameters, which are related to the time sequence of occurrence of errors.

3.5 *Cumulative Distributions*

Figures 3a, b, c, and d show weighted cumulative distributions for character error rate and lost character rate of calls. The unweighted distributions for all calls are shown for comparison.

The curves in Figs. 3a, b, c, and d have been left unsmoothed to reflect the actual results observed. Large discontinuities sometimes occur when the number of character errors or lost characters changes from 0 to 1, 1 to 2, etc.

The weighted results of average character error rates (Figs. 3a and b) for all calls are closer to the results for the short mileage band than

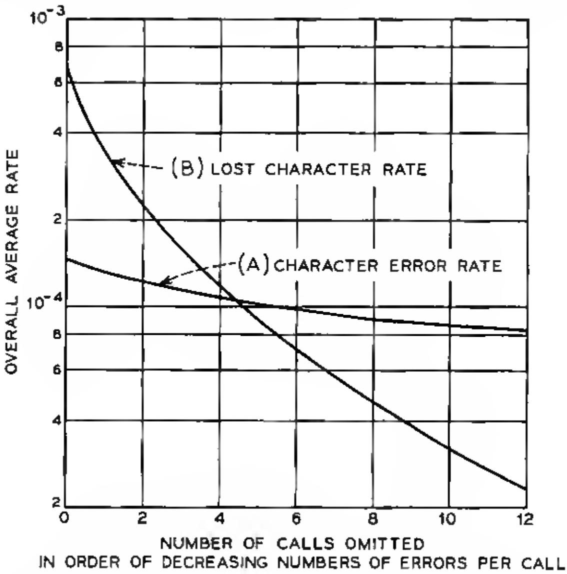

Fig. 2—Sensitivity of overall average rates to omission of worst calls.

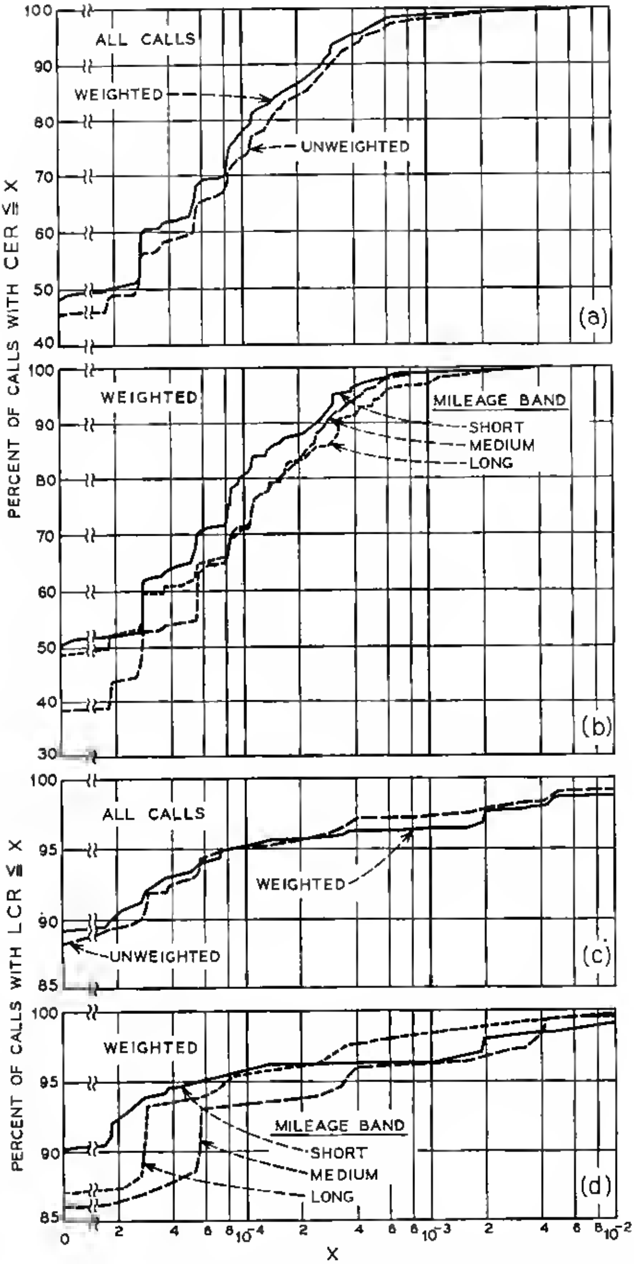

Fig. 3a—Cumulative distribution of average character error rate (CER)$_a$ of call.
Fig. 3b—Weighted cumulative distribution of average character error rate (CER)$_a$ of call.
Fig. 3c—Cumulative distribution of average lost character rate (LCR)$_a$ of call.
Fig. 3d—Weighted cumulative distribution of average lost character rate (LCR)$_a$ of call.

the unweighted data because this band has more toll traffic (85 percent of the total) and therefore a larger weighting factor than the other bands. However, application of weighting does not change the results markedly. At error rates above 10^{-4}, the results tend to get poorer as the mileage increases but not to a large degree.

Lost characters (Figs. 3c and d) were observed on relatively few calls; close to 90 percent of all calls have no lost characters. Their occurrence appears to be unrelated to mileage band. All distributions have long tails to the right because of the occurrence of a few long bursts of lost characters with rates per call greater than 10^{-2}.

Some of the results shown by the distributions are listed in Table III.

The weighted statistics show almost 50 percent of calls to be error free and almost 90 percent free of lost characters. About 78 percent of the calls have error rates equal to or less than 10^{-4}, and 95 percent of the calls have lost character rates equal to or less than 10^{-4}. About 81 percent have error rates better than the average for all calls; the corresponding number for lost characters is 96 percent.

Table III includes 90 percent confidence intervals for most of the parameters ($\hat{p} \pm \Delta$), meaning that there is 90 percent confidence that the true value of the parameter (p) for the entire population of switched network toll calls from which the survey sample was selected lies within the stated interval.*

The confidence interval for the average of lost character rate of calls (LCR) is very wide. This is because of the very large variance of LCR, the values of LCR ranging from 0 (89 percent of the calls) to 0.13. Ten calls had values of LCR greater than 0.01; the large departure from the average of 9.5×10^{-4} contributes to the large variance and wide confidence interval. The poor confidence indicated for this parameter is consistent with the sensitivity of the overall average LCR for all calls to the omission of a few bad calls, as noted in Section 3.3. It is concluded that average LCR for this survey is not a very significant parameter.

All other confidence intervals in Table III are reasonably narrow.

IV. BURST PROPERTIES

The distribution in time of character errors is far from uniform, nor is it Poisson (equal probability, overall average character error

* Stating this more precisely, if another survey were conducted, there would be 0.9 probability that the confidence interval for that survey, $\hat{p} \pm \Delta$, would contain p. Intervals were computed as described in Ref. 1. The formulas are derived for a normal distribution; the confidence intervals herein reported are only approximate because of departures of the distribution from normal.

TABLE III—ERROR STATISTICS OF CALLS

(A) Average Character Error Rate (CER) of Calls

	Weighted	Unweighted	Mileage Band (Weighted) Short	Medium	Long
Average CER*	$1.12 \pm 0.31 \times 10^{-4}$	1.46×10^{-4}	$1.03 \pm 0.39 \times 10^{-4}$	$1.22 \pm 0.62 \times 10^{-4}$	$1.56 \pm 0.56 \times 10^{-4}$
Percent of Calls with CER = 0	48.3 ± 8.5	45.8	50.5 ± 10.8	38.5 ± 7.5	48.9 ± 11.4
Percent of Calls with CER $\leqq 10^{-4}$	77.6 ± 5.8	73.4	80.1 ± 5.6	71.2 ± 8.5	70.8 ± 13.3
Percent of Calls with CER $\leqq$ Average	81.0	80.5	80.8	77.0	80.7

(B) Average Lost Character Rate (LCR) of Calls

	Weighted	Unweighted	Short	Medium	Long
Average LCR	$9.53 \pm 12.6 \times 10^{-4}$	6.8×10^{-4}	$1.16 \pm 1.78 \times 10^{-3}$	$2.82 \pm 2.53 \times 10^{-4}$	$4.91 \pm 7.56 \times 10^{-4}$
Percent of Calls with LCR = 0	89.3 ± 7.1	88.3	90.3 ± 8.0	86.0 ± 10.0	87.0 ± 6.6
Percent of Calls with LCR $\leqq 10^{-4}$	95.0 ± 3.5	95.1	95.4 ± 4.4	93.0 ± 4.5	95.4 ± 2.1
Percent of Calls with LCR $\leqq$ Average	96.4	97.2	96.0	94.2	97.8

* Note that $\pm$ values represent 90 percent confidence intervals.

rate, that any single character is received in error, regardless of the previous history). Instead, there is considerable bunching of character errors; once an error occurs, the probability of another error in any of the next several characters is much higher than the overall average character error rate.

These burst properties are shown in Figs. 4a and b. Burst density is here defined as the character error rate in an interval of L characters following a character error. No new interval is observed until the previous one is completed. Each interval (rather than each call) is weighted equally in computing the average burst density. This analysis of burst properties differs from that used in the analysis of high-speed voiceband error statistics.[2] Figure 4a is a plot of average burst density for all intervals on all calls in the survey, versus L. The probability that the character following an error is also received in error is about 0.45 for all calls, compared to 0.000146—the overall average character error rate. Complete subsiding of burst properties for interval lengths greater than a specific value of L would result in a plot approximately as shown (for $L = 10$), starting with a slope of unity (on a log-log plot). The actual plot for all calls settles into a straight line with slope less than unity beyond $L \cong 10$, indicating that there is little correlation of errors beyond this value, although there is no clear-cut maximum value of L at which burst properties subside completely.

Plots of average burst density for the three mileage bands separately are almost indistinguishable from that for all calls, indicating that burst properties are independent of distance.

The standard deviation of the distribution of burst densities is plotted versus L in Fig. 4b. The value for all calls decreases monotonically as L increases. The standard deviation is greater than the average (Fig. 4a) for all values of L; clearly, the distribution departs considerably from normal. Again, the plots for the separate mileage bands differ little from that for all calls.

V. ERROR-FREE INTERVALS

An error-free interval (EFI) occurs when one or more characters are received correctly between character errors. Its length (L) is the number of correctly received characters between the two character errors. Statistics of the duration of EFIs are of interest to users whose applications include error control.

5.1 *Interpretation of EFI at Start and End of Call*

The lengths of the EFIs preceding the first and following the last

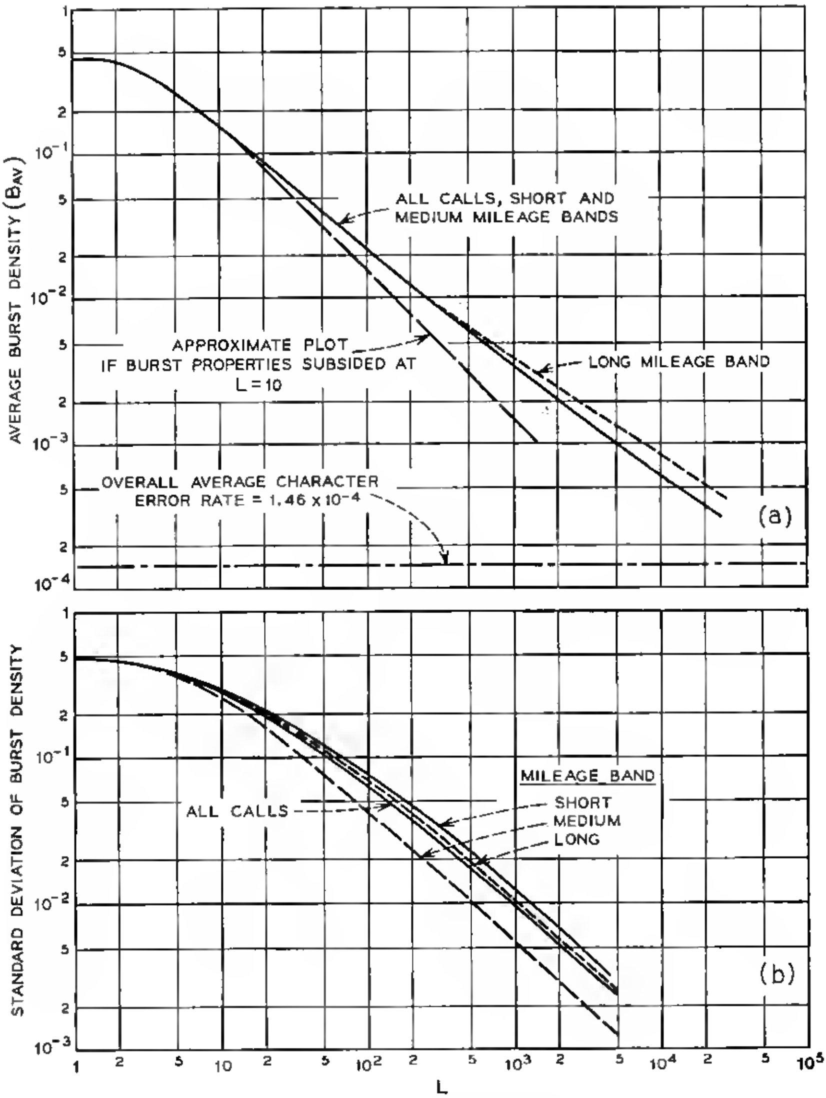

Fig. 4a—Average burst density versus L. Burst density (B) = character error rate in interval of length L (characters) following a character error.

Fig. 4b—Standard deviation of burst density versus L.

character errors in a call are unknown. These intervals are called end intervals (EIs). Since most calls have few if any errors, there are nearly as many EIs as complete error-free intervals.

An EI is part of an error-free interval which is at least as long as the EI. This information was used in estimating the probability, $P(\ell)$, that the length of an error-free interval is at least ℓ characters long, using the iterative procedure* of equation (1),

$$P(\ell) = \text{Prob } \{L \geqq \ell \mid L \geqq \ell - a\} P(\ell - a). \tag{1}$$

The procedure started with $P(0) = 1$. Each step required only the estimate of Prob $\{L \geqq \ell \mid L \geqq \ell - a\}$. This estimate equals the fraction of error-free intervals or EIs at least $\ell - a$ characters long which were also at least ℓ characters long. EIs shorter than ℓ were not used for this estimate.

5.2 *Cumulative Distribution of EFI Length*

This distribution is plotted in Fig. 5a. A Poisson distribution is shown for comparison. The difference between the survey and the Poisson curve shows the tendency of character errors to occur in bursts. For example, the survey data for all calls show that both short and long EFIs are more probable than for a Poisson distribution. For example, 19.5 percent of the survey EFIs are shorter than 100 characters compared to 1.45 percent for Poisson, and 48.6 percent of the survey EFIs are equal to or larger than 10,000 characters compared to 23.2 percent for Poisson. The latter comparison indicates that more error-free messages of long duration will result from the survey distribution than from Poisson. Equally spaced errors at the average error rate would result in all intervals having lengths of 6850 characters (Fig. 5a).

5.3 *Lost-Character-Free Intervals*

A lost-character-free interval (LFI) is the interval between lost characters; it is the corresponding parameter to EFI for lost characters. Its length is the number of received characters between the two lost characters.

The procedure described in Section 5.1 was followed for nonterminating lost character intervals. The cumulative distribution of LFI lengths is plotted in Fig. 5b. The plot shows that most LFIs tend to be much longer than EFIs (Fig. 5a); 95 percent exceed 10,000 characters for all calls. This seems inconsistent with the higher overall average lost character rate, almost five times the character error rate

* This is the same as the product limit method used in Ref. 2.

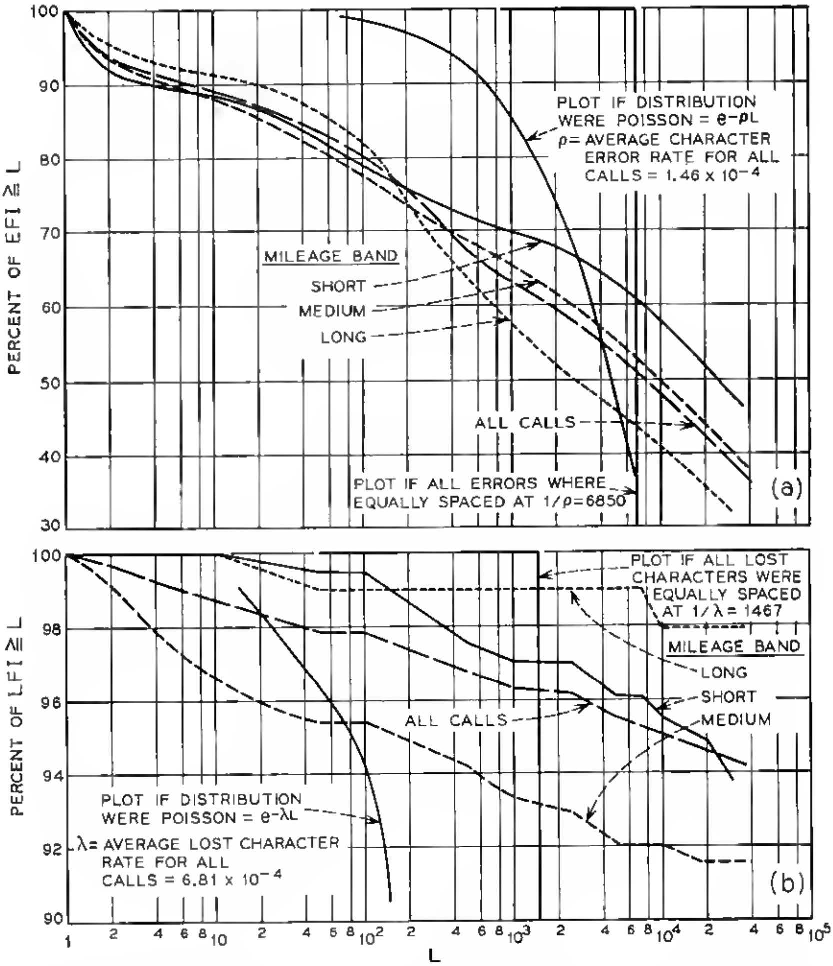

Fig. 5a—Cumulative distribution of character error-free interval (EFI) lengths (L characters).
Fig. 5b—Cumulative distribution of lost-character-free interval (LFI) lengths (L characters).

(Table II), but is explained by the occurrence of lost characters in long, continuous bursts.

VI. BLOCK ERROR STATISTICS

Error statistics in terms of blocks of many characters are useful in describing to a customer the accuracy that he might expect in transmission of messages of fixed length that he normally sends. For ex-

ample, the customer may be interested in what percent of his messages will be error free, or have no more than one error, etc. In addition, block error statistics are useful in evaluating performance of error control systems.[4,5]

Figures 6a and 6b give block error statistics for the survey data. These are obtained by dividing each call into consecutive blocks of equal duration (L characters). To obtain statistics independent of an arbitrary starting point, the division is made for 10 different starting points (phases) of the first block: 0, 1, 2, etc., tenths of a block from the beginning of the call. The statistics are then averaged for all 10 phases. Figure 6a shows average block error rate plotted against block length for various minimum numbers of character errors per block. Plots for uniform distribution and Poisson distribution of character errors are included for blocks with one or more errors.

Figure 6b shows cumulative distributions of block errors: probability of a block having more than E character errors versus E, for various block lengths.

The distribution for Poisson distributed errors is shown for blocks of 1000 characters. The bunching of errors in the survey data is shown by the comparison of the appropriate curves. The Poisson distribution has less probability of error-free blocks (0.864 versus 0.964) and greater probability of less than several errors (0.9999 versus 0.9928 for five errors); that is, the Poisson curve is much steeper than the survey curve.

VII. CARRIER FAILURES

Figure 7 shows a cumulative distribution of durations of carrier failures. These curves resemble the distributions of lost character rates per call (Figs. 3c and d). A large percentage of carrier failures have durations of one character or less (90.4 percent for all calls). Nevertheless, several long duration carrier failures (0.4 percent longer than 1000 characters) contain the bulk of the lost characters. The crisscrossing plots for the three mileage bands indicate the random occurrence of carrier failures and the lack of dependence on length of connection.

VIII. CONCLUSIONS

Several parameters describing error performance of low-speed, start-stop, data transmission over the Bell System switched telecommunications network have been presented. These include distributions

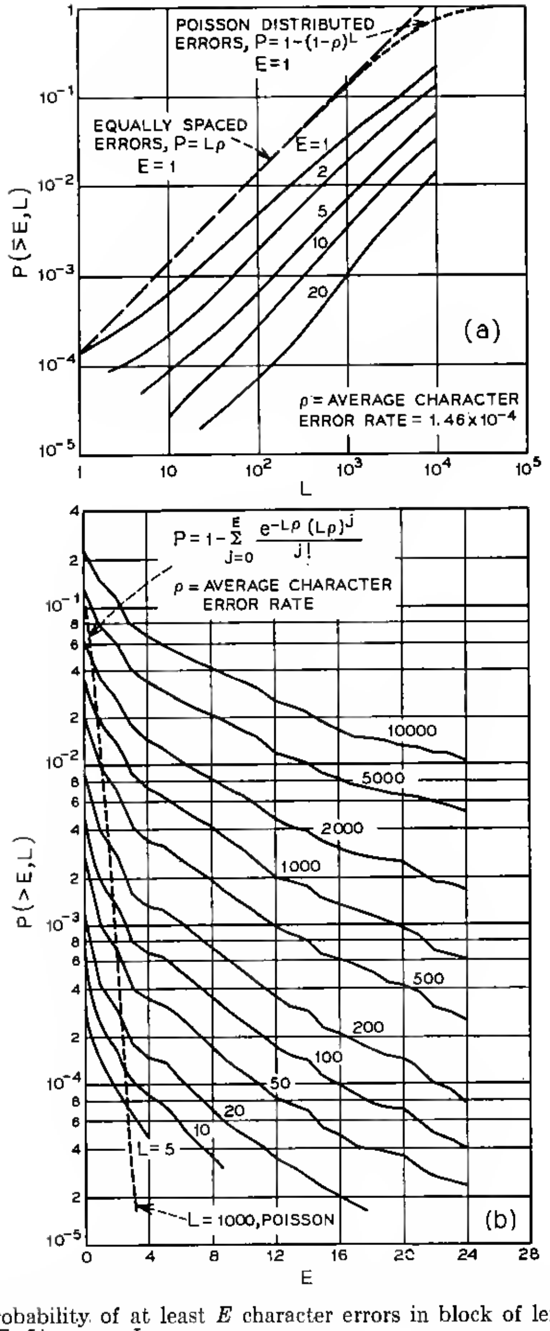

Fig. 6a—Probability of at least E character errors in block of length L (characters). $P(\geqq E, L)$ versus L.

Fig. 6b—Probability of more than E character errors in block of length L (characters). $P(>E, L)$ versus E.

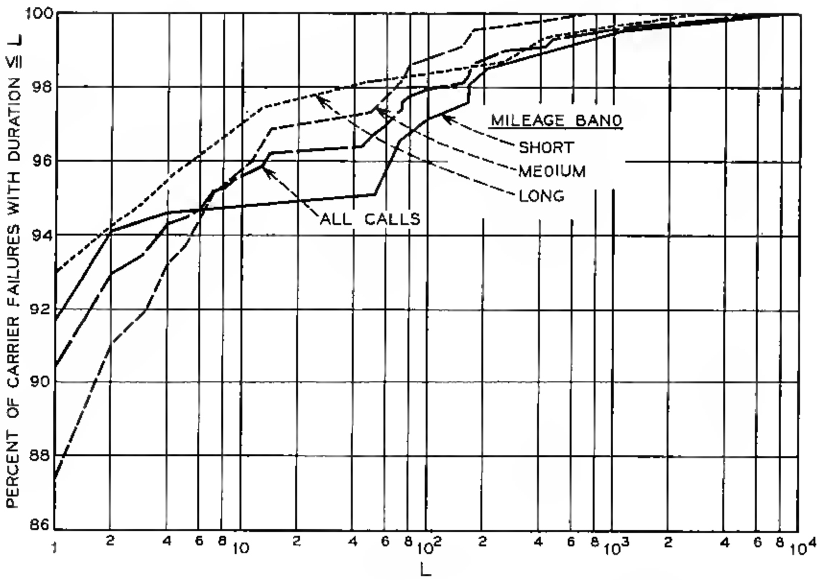

Fig. 7—Cumulative distribution of duration of carrier failures.

of averages per call of character error and lost character rates, and other parameters that relate to distributions in time of erroneous and missing characters. The intent has been to characterize accuracy of transmission from the sending data set input to the receiving device display by a teletypewriter or storage in a computer.

There have been no previous low-speed data surveys with which to compare the results of this one. Some comparisons with the results of the voiceband data measurements[2] are possible. The distribution of durations of carrier failures during low-speed data tests (Fig. 7) tails off to considerably longer periods than for the voiceband data tests (Ref. 2, Fig. 10). During low-speed tests, two percent of all calls had carrier failures longer than 7.5 seconds. The corresponding period for the voiceband tests was less than 0.3 second. The relatively long carrier failures and the resulting high lost character rate in the low-speed tests are believed to be fortuitous.

A comparison of error performance requires estimating the bit error rate at low speed, since this has not been computed. A rough approximation of the bit error rate is one-tenth of the character error rate since a character contains 10 bits. Isolated character errors have one bit or more in error, but this factor is compensated by the occurrence

of many character errors in OSS periods with no bit errors. Thus the low-speed result of 78 percent of calls with a character error rate of 10^{-4} or less may be compared with about 88 percent, for the 2000-b/s data set, and 84 percent, for the 1200-b/s data set, with a bit error rate of 10^{-5} or less (including bit errors during carrier failures), both greater than the percent for low speed. However, the overall average error rates for the voiceband data are about 1.9×10^{-5} (2000 b/s) and 6.6×10^{-5} (1200 b/s), greater than the low-speed value (approximately 1.46×10^{-5}). This apparent contradiction results from long tails on the voiceband distributions—a few calls with high error rates.

The results of the low-speed survey give much needed information on the error performance that may be expected on switched network calls. It is expected that further analysis of the statistics will give insight to the causes of errors which will suggest approaches to improve error performance.

IX. ACKNOWLEDGMENTS

Miss A. C. Weingartner and F. Brophy prepared the computer programs used to compact the data and determine the error statistics; the former also gave important assistance in the preparation of this article. Programs to separate the low-speed test results from the other records on the Holmdel computer magnetic tape were written by Mrs. K. D. Malleus. Editing of the preliminary records and other clerical tasks were performed by members of the Holmdel clerical pool, in particular, Mrs. S. D. Claus. Implementation and maintenance of test circuits was performed mostly by H. Hagen, who also devoted considerable time to the collection of data, as did many others.

REFERENCES

1. Duffy, F. P., and Thatcher, T. W., Jr., "1969–70 Connection Survey: Analog Transmission Performance on the Switched Telecommunications Network," B.S.T.J., this issue, pp. 1311–1347.
2. Balkovic, M. D., Klancer, H. W., Klare, S. W., and McGruther, W. G., "1969–70 Connection Survey: High-Speed Voiceband Data Transmission Performance on the Switched Telecommunications Network," B.S.T.J., this issue, pp. 1349–1384.
3. Model 37 Regenerator—Circuit Description 8383 WD, Issue 3, Teletype Corporation, Skokie, Illinois.
4. Elliott, E. O., "A Model of the Switched Telephone Network for Data Communications," B.S.T.J., *44*, No. 1 (January 1965), pp. 85–109.
5. Elliott, E. O., "Estimates of Error Rates for Codes on Burst-Noise Channels," B.S.T.J., *42*, No. 5 (September 1963), pp. 1977–1998.